Network Outlier Detection in a single large graph

Zhihao Chen

What is a network outlier?

Outlier: An observation which diverges from the overall pattern

Social Media - block structure

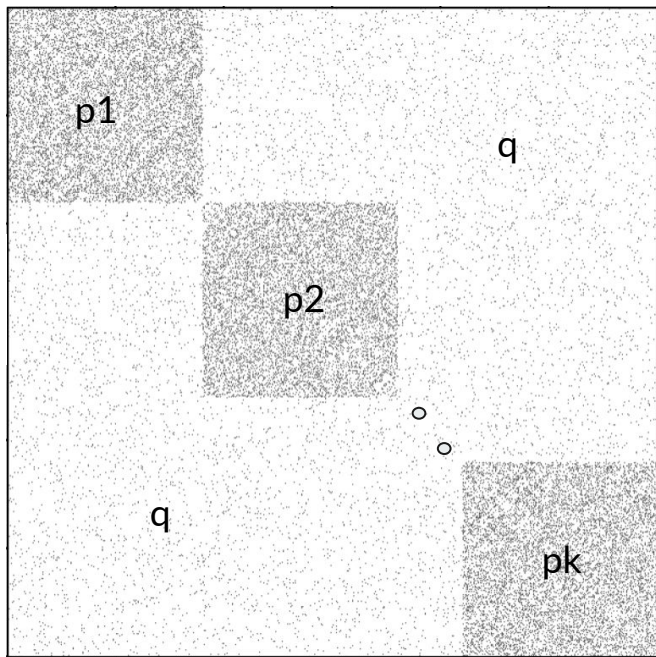Network Outlier: A node that doesn't fit the block structure

Step 1: Find the blocks

Spectral clustering

1. Laplacian matrix = Degree matrix - Adjacency matrix
2. K-means on the eigenvectors for the k smallest eigenvalues

Step 2: Outlier score

$$Bern(p_i) \quad\quad\quad Bern(q)$$

| 1 0 1 ... 1 | 0 1 0 0 0 .... 0 1 0 |

$$n_i \quad\quad\quad\quad N - n_i$$

Obtain the residual matrix: $R = A - \hat{A}$

Outlier score for node i $= \sum_j R_{ij}^2$

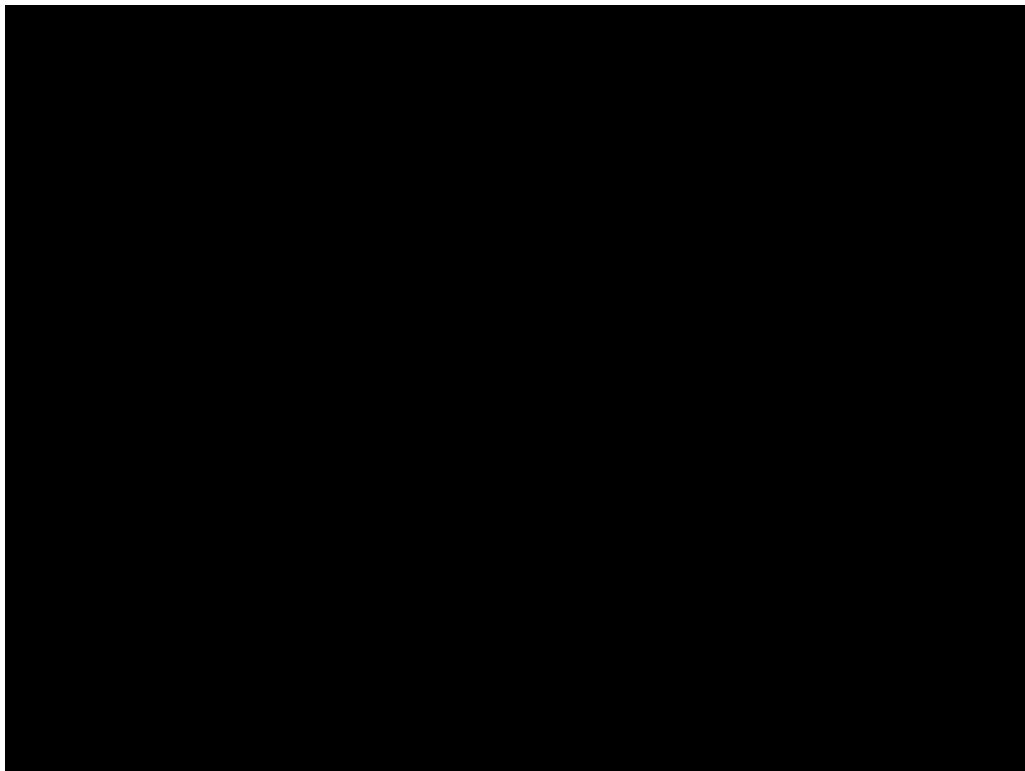Outlier score has an approximately normal distribution
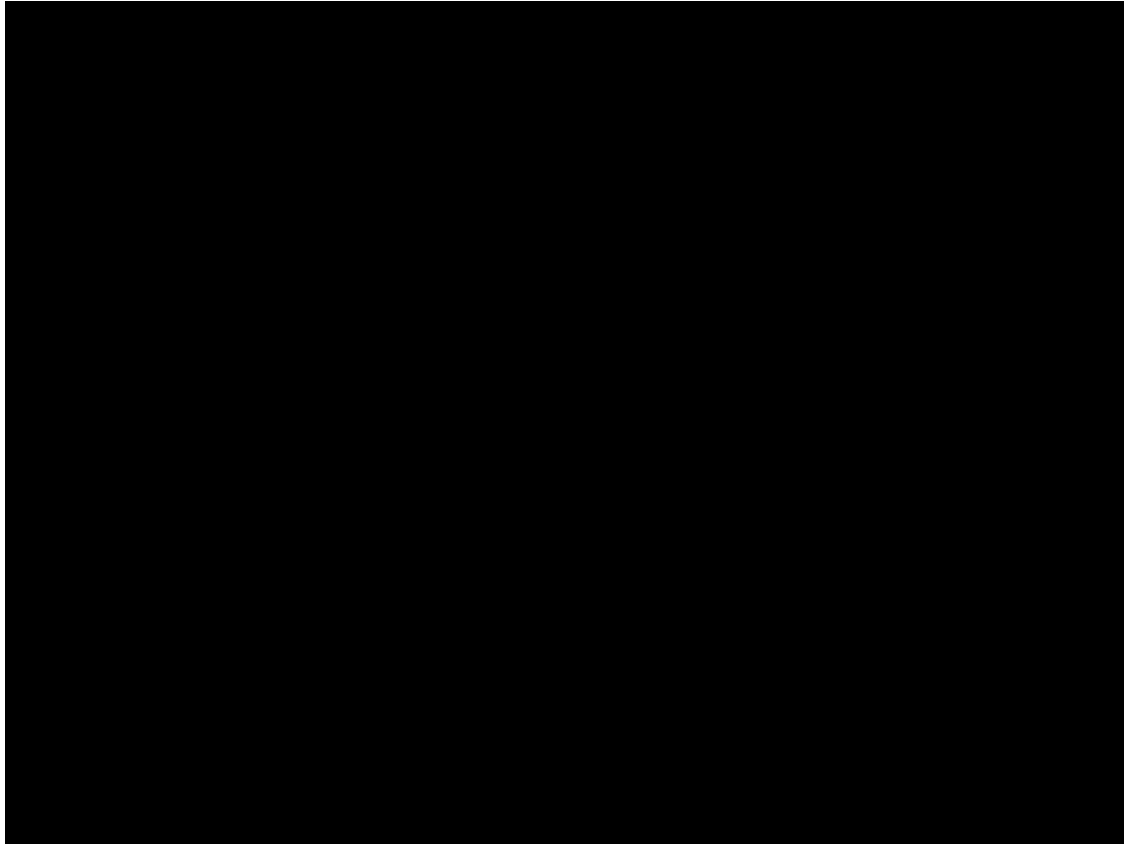
$$\hat{A} =$$

Simulations

Network:

3 communities, each 100 nodes

p : q = 1 : 0.2

overall degree = 50

Add edges to 20 nodes:

Power Comparison

Jing Lei's method:

Our method:

Summary

A test for outliers that don't fit the block structure - not necessarily have the largest degree.

Further research on the performance of this test when the spectral clustering is not working well.

Thank you!

Next: Zesheng - Network Outlier Detection in many small graphs